DOCUMENT RESUME

ED 187 558 SE 030 944

AUTHOR Farmer, Walter A.; And Others
TITLE Test Competency as Related to Classroom Performance.
PUB DATE May 77
NOTE 15p.; Paper presented at the Annual Meeting of the New England Educational Research Association (Manchester, NH, May 5, 1977). Contains occasional light and broken type.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Behavior Rating Scales; *Competency Based Teacher Education; *Educational Research; Higher Education; *Mathematics Teachers; Personality Measures; Research; *Science Teachers; Student Teachers; Student Teaching; *Teacher Behavior; *Teacher Education; Tests

ABSTRACT

This study of student teachers in high school mathematics and science classrooms compares their teaching behaviors and their scores from three pencil-and-paper tests. Data analysis indicates a relative independence between the ratings of teacher classroom behavior and scores on the pencil-and-paper tests. There was little consistent relationship between how well student teachers could judge what they ought to do in a classroom situation and what they actually did. Apparently, personality dimensions, as measured by the Myers-Briggs, tell little about what to expect from a person in the classroom. (Author/NK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED187558

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Walter A. Farmer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TEST COMPETENCY as RELATED to CLASSROOM PERFORMANCE

by

Walter A. Farmer, Margaret A. Farrell and Richard M. Clark
State University of New York at Albany

A research paper presented by Walter A. Farmer at the Annual Meeting of the New England Educational Research Association, May 5, 1977, Manchester, New Hampshire.

SE 030 944

Test Competency as Related to Classroom Competency

Walter A. Farmer, Margaret A. Farrell and Richard M. Clark

State University of New York at Albany

In a previous study Farrell, Farmer, and Clark (1976) reported on the teaching behaviors of 100 student teachers in high school math and science classes. In that study a list of teaching behaviors to be assessed was developed from a model developed by Farrell and Farmer. Behaviors were factor analyzed and the factor structure was compared to the logical model which had been developed.

In the present study the same set of teaching behaviors are considered along with several scores derived from paper and pencil tests. In the process of teacher education, and especially in a competency based model, questions arise as to how to assess teacher skills. For example, to what degree does a score on a test relate to what student teachers actually do in classroom situations? If a factor analysis is performed, to what extent will personality test measures and other paper/pencil measures load with behavioral measures? To carry out this study, four assessment devices were employed. Each of these devices will be briefly described.

Teacher Classroom Behavior Rating

Based on the work of Gagne, Ausubel, Piaget, and Bruner, a model was developed of seventy-four teacher competencies which could be observed in a classroom situation. These behaviors were organized into a model previously presented (Farrell, Farmer, and Clark, 1976). Some subject-specific behaviors in the original model were applicable to either (not both) math or science teachers, although most behaviors were applicable to both groups. In this study the 57 teacher behaviors which were appropriate for both math and

science teachers made up the Teacher Classroom Behavior Rating. Examples of these behaviors are seen in the Tables in which factor analyses are reported.

Data on teaching behaviors were gathered over a period of five semesters from spring 1973 through spring 1975 on the classroom performance of 122 enrollees, of which 100 successfully completed the program. Each student teacher was observed by either Farrell or Farmer a minimum of eight-45 minute teaching periods during the ten weeks of student teaching. A similar number and kind of written observation were also compiled by the public school cooperating teacher to whom each student was assigned. Summative evaluation based on the data resulted in an assessment of one of three kinds: below criterion, meets criterion, or exceeds criterion. (Exceeds criterion was achieved by a very small sample of students.) "Criterion" was defined as the performance of the competency effectively and over at least the final two-three weeks of student teaching. In order to maximize inter-observer reliability in data-gathering, the above researchers frequently observed the same class session, privately recorded relevant data, and then cross-checked the written record. Likewise, these observers regularly read data collected by each other and discussed the kind of inferences to be drawn as related to the AMST instrument.

One hundred students successfully completed the program. These students were to have taken three different tests. Because some students missed testing sessions, the final N for this study consists of 41 math teachers and 31 science teachers for a total of 72 subjects.

AMST Test

The AMST Test was derived from the behaviors which were to be determined from classroom observations. Section 1 of the test employed a multiple choice format in which the stem described a specific event that might take place in the class. The student was then asked to choose one of five optional answers.

Each item had an answer keyed as worth three points, one or more items worth two points, and the other responses worth only one point. Part one consists of nine items.

A sample item and the scoring for the item is presented below:

1. After directing the class on how to solve a certain type of problem all students, with the exception of Jane, show evidence of satisfactory comprehension. Jane (a slow student) asks T questions about this type of problem which indicates that she is completely confused. Teacher should:

Score

-1 a) Tell Jane not to worry about it, accepting the fact that Jane cannot do the work.

1 b) Have the class wait while T goes through another example for Jane's benefit.

-2 c) After school provide for individualized instruction for Jane by going over the problem step by step.

1 d) Encourage Jane by telling her that if she keeps trying she will learn it.

3 e) After school ask Jane a series of diagnostic questions to locate the point at which her understanding ceased.

Parts 2 and 3 of the test are in brief essay format. Again, a classroom situation is described and the teacher is asked to indicate how he or she would respond to the situation. The items in part 2 concern how the teacher would deal with problems of feedback and motivation and part three with problems related to subject matter in math and science.

High scores on all three parts of the test were assumed to indicate that students had an intellectual understanding of teacher classroom behavior which would be scored at criterion level on the final evaluation instrument for student

teaching. If high scoring students were able to translate their intellectual understandings into action in their teaching behaviors, then they would display many of the teaching behaviors rated.

Myers-Briggs Type Indicator

The Myers-Briggs Type Indicator is a standardized test designed to classify people into four personality dimensions and then describe the strength of the personality dimension. For example, the test score classifies an individual as extroverted or introverted and then indicates the strength of introversion or extroversion. The test makers prefer that analyses be done separately for introversion and extroversion rather than treating the scores as a continuous variable ranging from extreme introversion to extreme extroversion. They do provide, however, a procedure for deriving a continuous score, and this procedure was used in this study.

The means and standard deviations of Myers-Briggs scores are reported in Table 5. Also included are brief definitions of the four personality dimensions. As can be seen, the general tendency of the group is to be introverted rather than extroverted, feeling rather than thinking, and uses judgment rather than perception. The group is about evenly split on the dimension of sensing or intuition.

Mathematics and Science Teaching Inventories - Teacher Perceptions

The Mathematics and Science Teaching Inventories are designed to measure the perceptions of student teachers from a low score of rigid, fixed, arbitrary behavior to a high of open, flexible, inter-active behavior. The instrument is divided into four parts, under the assumption that different perceptions are possible in different areas of the teacher's responsibility. The four parts of the test are:

(1) perception of student and teacher roles

(2) perception of the appropriate use of textbooks and other instructional material

(3) perception of the appropriate means of designing and using tests

(4) perception of appropriate strategies of teaching.

Each section is composed of a set of declarative sentences which a student rates on a four-point scale from "strongly agree" to "strongly disagree." No neutral point is included on the scale.

Analysis of Data

Data collected from the four psychometrics described were analyzed in the following ways:

(1) A principle component factor analysis was performed on the set of variables including the 57 teaching behaviors, 3 AMST scores, 4 Myers-Briggs scores, and 4 Teacher Perception scores, the sex of teacher, and the major of teacher (math or science). It is recognized that too many variables were employed for the number of subjects, but we wished to see the degree to which test variables would factor in with behavioral variables. The behaviors in Factors 1 and 2 are presented in Tables 1 and 2. Only behaviors with factor loadings over 5.00 are presented. The proportion of total variance accounted for by factor one is .305 and the cumulative portion of total variance for factors 1 and 2 is .377. The first six factors had eigenvalues larger than 2.00, accounting for .534 of the total variance. However, relatively few variables accounted for factors 3-6.

(2) Means, standard deviations, and correlations of the sub-parts of each of the three pencil-paper instruments are presented. These data are presented in Tables 3, 4 and 5. As can be seen, the sub-parts of the Myers-Briggs and the AMST Test are relatively independent, although

some small but significant sub-part correlations are present for the AMST. Sub-parts of the Perception test are significantly interrelated, but the correlations are not high.

(3) A discriminate function-analysis was performed to contrast the responses of 41 math teachers from those of the 31 science teachers. (See Table 6)

Discussion

Probably the most interesting finding of this study is the relative independence of the ratings of teacher classroom behavior from the scores on the paper and pencil tests. Factor analysis produced two major factors which implicated more than half of the total set of variables. Yet none of the 13 variables derived from the paper and pencil tests had significant loadings in the first two factors. Apparently there was little consistent relationship between how well student teachers could judge what they ought to do in a classroom situation and what they actually did. Apparently also, personality dimensions as measured by the Myers-Briggs tell us little about what to expect from a person in the classroom.

For those concerned with competency based teacher education, the results of this study should raise a series of questions. We assume that the most valid measures of competency employed in this study were derived from the behaviors displayed by teachers in the classroom. If this assumption is accepted, then the various paper and pencil measures employed in this study are essentially invalid measures of teacher competency although many of the measures were derived from classroom examples. Many competency based programs are based on the assumption that answers to paper-pencil measures can indeed certify competence, and further assume that such competence will be displayed in classroom situations. These assumptions do not appear to be warranted based on our data.

One reason we may have difficulty in finding more significant relationships may relate to the sample employed in this study. First, highly selective

admission and retention policies of the University serve to reduce variance in those entering the AMST program in their senior year. Second, all of the data analyzed were obtained from those judged to have succeeded in the program. Different relationships might have been found if data were available from a wider range of students and from students who failed to complete the program.

Also of interest are the clusters of behaviors within the first two factors. In general, the first factor seems to represent behaviors that are relatively concrete and specific. If a student teacher were told to "work on" the behaviors in the first factor, he or she might have a relatively clear idea as to what would be needed. Many factor two variables seem to involve more subtle abilities. Just what does one do to arouse enthusiasm, maintain control, capture attention, and meet individual needs? Somehow, the problem seems different from such factor one behaviors as constructing good tests, trying out demonstrations, and keeping the heat at a proper level. In general, some of the variation for factor 1 items came from students exceeding the criterion set rather than failing to reach the criterion. Factor 2 behaviors were generally harder to attain as evidenced by the number of students who didn't attain them.

The results of the discriminate function analysis to contrast science and math student teachers are not especially surprising. Demonstrations, for example, may be used in either math or science, but it is not surprising to find that more science teachers are likely to be rated successful in this competency. Questions could certainly be raised as to the degree that teacher competencies are global or specific. To what degree do science and math teachers really need the same competencies? To what degree do English and science teachers? Obviously the subject matter that needs to be known will differ from one field to another, but should we look for other differences? Is it more important that the English teacher be skillful in leading group discussions than that the math teacher have this ability?

We are far from having either an adequate logical or empirical base for competency based teacher education. Hopefully, this study is one step in the proper direction.

Table 1

BEHAVIORS AND FACTOR LOADINGS FOR FACTOR 1

Behavior	Factor Loading
Maintained a physical environment conducive to learning (e.g., lights, heat, ventilation).	.901
Constructed tests to be consistent with emphasis of instruction.	.856
Used more than one mode during instructional period.	.827
Demonstrated clerical efficiency with respect to the paperwork of the teacher.	.827
Provided a model for correct performance of skills to be learned (to enable students to get feedback on their efforts).	.818
Matched test items to instructional objectives.	.817
Managed and utilized the physical organization of the room, equipment and materials.	.759
Used out-of-class time to help individual students.	.723
Related present instruction to future lessons.	.712
Made formal oral presentations (short or long duration).	.707
Tried out demonstrations before presenting them in class.	.706
Employed means other than written tests to gather evidence of learning.	.695
Gave immediate feedback to individuals and groups.	.672
Produced legible and accurate written communications (board, overhead, handouts).	.670
Provided for and supervised student practice.	.650
Employed appropriate volume, modulation, pitch, and reasonable clarity of voice.	.634
Organized instructional materials in a systematic way.	.633
Maintained eye contact with a wide sampling of students.	.632
Demonstrated regard for the worth of individual personalities of students.	.609
Sequenced subject matter in ways meaningful to students.	.605
Designed individualized instructional materials based on varying objectives and learning styles.	.586
Used instructional strategies which matched stated objectives.	.583
Designed assignments which reinforced the concepts, rules, or skills taught in a previous class session.	.509

Table 2

BEHAVIORS AND FACTOR LOADINGS FOR FACTOR 2

Behavior	Factor Loading
Used positive techniques to prevent and/or control student behavior.	.765
Generated enthusiasm for the subject among students.	.678
Provided a classroom atmosphere which encouraged students to demonstrate regard for the rights and property of others.	.655
Developed presentations which captured student attention at the start of lessons.	.610
Used instructional strategies consistent with the intellectual developmental stages of the student.	.608
Arranged instructional modes so that student attention was maintained throughout the lessons.	.589
Asked thought-provoking and/or open-ended questions.	.582
Included those objectives designed to effect the attainment of intellectual skills beyond the level of recall type problems.	.576
Designed differentiated assignments based on student interests and levels of previous learning.	.574
Designed assignments to lead to the discovery of concepts, rules, or skills to be taught in subsequent class session.	.558
Collected feedback frequently from a broad sampling of students.	.541
Included objectives designed to effect positive attitudinal changes toward subject matter.	.533

Table 3

SUMMARY DATA OF AMST TEST

Mean, Standard Deviation, and Range of AMST Test

		Mean	S.D.	Range
(Management)	Part 1	20.1	2.7	15-26
(Feedback Motivation)	Part 2	9.4	1.4	6-14
Subject specific	Part 3	8.6	2.1	5-14

Intercorrelations of Sub-test Scores

	2	3
1	.224**	-.055
2		.277**

** sign. at .01 level

Table 4

SUMMARY DATA MATH-SCIENCE PERCEPTION TEST

Perception of	Mean	S.D.
1 Student and teacher role	25.7	2.38
2 Use of textbooks	28.5	2.36
3 Use of tests	25.6	2.80
4 Strategies of teaching	80.0	17.5

Intercorrelations of Test Sub-Parts**

	2	3	4
1	.274	.419	.369
2		.368	.215
3			.470

** all correlations are sig. at .01 level

Table 5

SUMMARY DATA ON MYERS-BRIGGS

Below 100	Above 100	Mean	S.D.
1 Extroversion	Introversion	107.4	26.8
2 Sensing	Intuition	101.9	28.8
3 Thinking	Feeling	110.3	24.1
4 Judgment	Perception	92.8	29.3

Intercorrelation of Sub-Test Scores

	2	3	4
1	-.275**	.001	-.05
2		-.053	.341**
3			.100

** sig. at .01 level

Definitions of Terms

Extroversion
Tends to focus perception and judgment on people and things.

Introversion
Tends to focus perception and judgment on concepts and ideas.

Sensing
Made aware of things primarily through use of one or another of his five senses.

Intuition
Made aware of things through indirect perception with emphasis on ideas and associations through the unconscious.

Thinking
Discriminates impersonally between the true and false.

Feeling
Discriminates personally between the valued and not-valued

Judgment
Relies primarily on a judging process in dealing with outside world.

Perception
Relies on a perceptual process in dealing with outside world.

Table 6

BEHAVIORS WHICH MOST CLEARLY DISCRIMINATED

BETWEEN MATH AND SCIENCE TEACHERS**

Behaviors	Mean Score	
	Math	Science
Utilized laboratory activities (student manipulation of materials).	1.44	2.06
Designed assignments to lead to the discovery of concepts, rules or skills to be taught in subsequent class sessions.	1.41	1.19
Made immediate use of feedback to pace instruction.	1.53	2.00
Provided for applications of subject matter based on the "here and now" interests of individual students.	1.58	1.93

** All differences sig. at .01 level